

Efficient classical simulations of quantum Fourier transforms and normalizer circuits over Abelian groups

Maarten Van den Nest*

Max-Planck-Institut für Quantenoptik,
Hans-Kopfermann-Straße 1, D-85748 Garching, Germany.

January 25, 2012

Abstract

The quantum Fourier transform (QFT) is sometimes said to be the source of various exponential quantum speed-ups. In this paper we introduce a class of quantum circuits which cannot outperform classical computers even though the QFT constitutes an essential component. More precisely, we consider *normalizer circuits*. A normalizer circuit over a finite Abelian group is any quantum circuit comprising the QFT over the group, gates which compute automorphisms and gates which realize quadratic functions on the group. We prove that all normalizer circuits have polynomial-time classical simulations. The proof uses algorithms for linear diophantine equation solving and the monomial matrix formalism introduced in our earlier work. We subsequently discuss several aspects of normalizer circuits. First we show that our result generalizes the Gottesman-Knill theorem. Furthermore we highlight connections to Shor's factoring algorithm and to the Abelian hidden subgroup problem in general. Finally we prove that quantum factoring cannot be realized as a normalizer circuit owing to its modular exponentiation subroutine.

1 Introduction

The quantum Fourier transform (QFT) is arguably the most important operation in the field of quantum algorithms. Indeed this transform is a principal component in several quintessential algorithms (cf. [1, 2] for reviews), most importantly Shor's factoring algorithm [3]. In fact it is sometimes said that the QFT is the crucial non-classical ingredient accounting for the (believed) speed-up of quantum factoring.

Whereas in quantum algorithm design the focus lies on identifying scenarios where usage of the QFT yields new quantum speed-ups, in this paper a complementary question is asked: under which conditions does a quantum algorithm, in spite of its use of the QFT, *fail* to outperform classical computation? Insight in this problem may shed light on the QFT's function in quantum computation and indicate avenues for new quantum algorithms. In addition, since quantum algorithms have provided new ways to relate Fourier transforms to various problems in e.g. number theory and group theory (the central example being the factoring problem) a deeper understanding of the QFT might also lead to better *classical* algorithms for these problems.

Whereas colloquially one often refers to “the” QFT as if it were a single operation, a different QFT is in fact associated with every finite group. The standard QFT is defined over the Abelian group of integers modulo 2^n . Throughout this work we focus on QFTs

*maarten.vandennest@mpq.mpg.de

over arbitrary finite Abelian groups. This is arguably the class of QFTs which has been most successfully employed in quantum algorithms [2].

We will introduce a class of quantum circuits, called here *normalizer circuits*. A normalizer circuit over an Abelian group G is any circuit containing the QFT over G and two other types of operations, called automorphism gates and quadratic phase gates, both of which have natural group-theoretic definitions. In our main result we prove that all polynomial-size normalizer circuits, acting on a natural class of inputs (called coset states) and followed by a standard basis measurement, can be simulated classically in polynomial time.

This result sharply contrasts with the QFT’s reputation of being classically intractable. In particular it highlights (complementary to previous works, cf. below) that the believed speed-up of e.g. Shor’s algorithm cannot straightforwardly be attributed to the mere presence of the QFT: a more subtle mechanism must be responsible for this. We will go into some detail analyzing the relation between normalizer circuits and quantum factoring. In particular we will prove that Shor’s algorithm cannot be realized by a normalizer circuit, showing that (non-surprisingly) our result does not yield an efficient classical factoring algorithm. The single component of Shor’s algorithm that is not covered by the normalizer gates formalism is modular exponentiation, which is sometimes (mistakenly, in our opinion) considered to be a simple “classical” step. In spite of this no-go result, we will highlight close connections between normalizer circuits and quantum factoring, and the Abelian hidden subgroup problem (HSP) in general¹. Interestingly, the quantum states occurring in the HSP algorithm and those generated by normalizer circuits turn out to be very similar and, all in all, we find it very difficult to give a concise intuitive reason that clarifies why normalizer circuits have efficient classical simulations whereas the factoring algorithm (probably) does not. In particular the usual superficial arguments that the power of quantum factoring comes from entanglement, interference or other “quantumness” measures, seem to apply equally well to normalizer circuits. Probably this indicates that such arguments have little merit and that much more delicate considerations are necessary.

Finally we remark that our results generalize the celebrated Gottesman-Knill theorem which regards classical simulations of Clifford circuits [6]. In the language of the present work, Clifford circuits correspond to normalizer circuits over the group $\mathbf{Z}_2 \times \dots \times \mathbf{Z}_2$.

Relation to previous work. To our knowledge, classical simulations of QFTs were previously investigated in three independent works; these results are similar to one another but different from ours. In Ref. [7] it was shown that the approximate QFT over \mathbf{Z}_{2^n} acting on a product input state and followed immediately by measurement in a product basis can be simulated classically in quasi-polynomial $O(n^{\log n})$ time. In Ref. [8] the same scenario was considered and a fully efficient classical simulation was achieved; more generally, efficient simulations were obtained for constant-depth circuits of bounded interaction range, interspersed with a constant number of approximate QFTs. In Ref. [9] efficient simulations were obtained for the “semi-classical” QFT acting on a class of entangled input states. The common ingredient in these works is their use of tensor contraction schemes which depend crucially on the *structure* of the quantum circuit. These methods are only efficient when the graph representing the topology of the circuit is sufficiently close to a tree graph as measured by the tree-width [11]. A common feature of such circuits is that they can only generate limited amounts of entanglement (of certain type) [12]. The fundamental distinction with the

¹Recall that the HSP encompasses the factoring problem as a special instance [4, 5].

present work is that our classical simulations are *independent* of the structure of the circuit. This comes at the cost of restricting the allowed operations, similar to the Gottesman-Knill theorem. Finally we mention that in [10] it was shown that the existence of an efficient classical algorithm to simulate the modular exponentiation circuit acting on arbitrary product input states and for arbitrary product measurement, would imply an efficient simulation of Shor's algorithm. This result is similar in spirit to our result that modular exponentiation is the only component of quantum factoring that cannot be realized by a normalizer circuit.

Organization of this paper. Readers who are mainly interested in the results rather than their proofs may focus on sections 2-5 where we introduce normalizer circuits, describe our main result and discuss its relation to Shor's algorithm and the HSP. In subsequent sections mathematical tools are developed and proofs of our main results are given. Considerable effort goes into laying out the necessary group theoretic concepts and their computational complexity, cf. section 6. In section 7 we discuss generalized Pauli and Clifford operators over Abelian groups. In section 8 we recall recently introduced techniques to describe and classically simulate a class of many body-states called M-states [24]. In section 9 we prove our main classical simulation result by combining the methods developed in sections 6-8. In section 10 we prove that Shor's factoring algorithm cannot be realized as a normalizer circuit. Finally in section 11 we give more examples of normalizer circuits.

2 Normalizer circuits

Throughout this paper the symbol G will represent a finite Abelian group given as a direct product of cyclic groups:

$$G = \mathbf{Z}_{d_1} \times \cdots \times \mathbf{Z}_{d_m} \quad (1)$$

where \mathbf{Z}_{d_i} denotes the additive group of integers modulo d_i . It is a fundamental result that every finite Abelian group is isomorphic to a decomposition of this form. Here we assume that G is presented *explicitly* in this way; this is a nontrivial constraint since computing this decomposition is generally a hard task. Every group element is a tuple $g = g_1 \cdots g_m$ where $g_i \in \mathbf{Z}_{d_i}$. Addition of group elements is performed componentwise where arithmetic in the i -th component is modulo d_i . The cardinality of G is denoted by \mathfrak{g} . For every i we introduce a Hilbert space \mathcal{H}_i with orthonormal basis $\{|x\rangle : x \in \mathbf{Z}_{d_i}\}$. Taking the tensor product of these m spaces we obtain a \mathfrak{g} -dimensional space \mathcal{H} . Defining $|g\rangle := |g_1\rangle \otimes \cdots \otimes |g_m\rangle$, the space \mathcal{H} has a tensor product basis where each basis vector is labeled by an element of G . This basis will be called the standard basis. Note that the above construction of the space \mathcal{H} is commonly used in the field of quantum algorithms (cf. e.g. [2]). We now define three classes of unitary operators on \mathcal{H} .

Quantum Fourier transforms. The quantum Fourier transform (QFT) over \mathbf{Z}_{d_i} is the following unitary operator on \mathcal{H}_i :

$$F_i = \frac{1}{\sqrt{d_i}} \sum \omega_i^{xy} |x\rangle \langle y|; \quad \omega_i = \exp(2\pi i/d_i) \quad (2)$$

where the sum is over all $x, y \in \mathbf{Z}_{d_i}$. The QFT over the entire group G is given by $F = F_1 \otimes \cdots \otimes F_m$. Any operator obtained by replacing a subset of operators F_i in this tensor product by identity operators will be called a partial QFT.

Automorphism gates. An automorphism α is an isomorphism from G to itself. The associated automorphism gate U_α maps $|g\rangle \rightarrow |\alpha(g)\rangle$. Since every automorphism is invertible, U_α is a unitary operator which acts a permutation on the standard basis.

Quadratic phase gates. A function B from $G \times G$ to the nonzero complex numbers is said to be bilinear if for every $g, h, x \in G$ one has

$$B(g + h, x) = B(g, x)B(h, x) \quad \text{and} \quad B(x, g + h) = B(x, g)B(x, h). \quad (3)$$

A function ξ from G to the nonzero complex numbers is said to be quadratic if there exists a bilinear function B such that for every $g, h \in G$ one has

$$\xi(g + h) = \xi(g)\xi(h)B(g, h). \quad (4)$$

Given a quadratic function ξ , the quadratic phase gate D_ξ is the diagonal operator mapping $|g\rangle \rightarrow \xi(g)|g\rangle$. In section 6.3 we will show that $\xi(g)$ is a complex phase for every $g \in G$. This implies that every quadratic phase gate is a unitary operator.

A unitary operator which is either a (partial) quantum Fourier transform or its inverse, an automorphism gate or a quadratic phase gate will generally be referred to as a normalizer gate; the origin of this nomenclature will be clarified in section 7. A quantum circuit composed entirely of normalizer gates is called a normalizer circuit over G . The size of a normalizer circuit is the number of normalizer gates of which it consists.

We further recall how computational cost is measured in the present context. A (classical or quantum) computation is called efficient if the number of elementary operations of which it consists scales polynomially in the logarithm of \mathfrak{g} . Since \mathfrak{g} coincides with the dimension of \mathcal{H} , the above definition is compatible with the standard concept of efficiency in quantum computation viz. the number of operations scales polynomially with the logarithm of the Hilbert space dimension. In this work we will subsequently consider families of normalizer circuits of size $\text{polylog}(\mathfrak{g})$, henceforth called polynomial size normalizer circuits.

It is not evident that each individual normalizer gate can itself be efficiently realized on a quantum computer i.e. as a quantum circuit composed of $\text{polylog}(\mathfrak{g})$ elementary gates. It is well known that (partial) quantum Fourier transforms over arbitrary finite Abelian groups can be realized as efficient quantum circuits, either approximately [13] or exactly [14]. In appendix A we show that all automorphism gates and quadratic phase gates can also be realized exactly as efficient quantum circuits.

3 Examples

We illustrate the definitions of normalizer gates with two examples viz. \mathbf{Z}_2^m and \mathbf{Z}_{2^m} . Both groups have $\mathfrak{g} = 2^m$ so the notion “efficient” is synonymous to “in $\text{poly}(m)$ time”. See also section 11 for further examples of normalizer gates.

3.1 \mathbf{Z}_2^m and Clifford circuits

The group \mathbf{Z}_2^m consists of all m -bit strings with addition modulo 2. Each Hilbert space \mathcal{H}_i is a single qubit with basis $\{|0\rangle, |1\rangle\}$ so the total Hilbert space \mathcal{H} is a system of m qubits. The standard basis is the usual computational basis. Applying definition (2) one finds that the QFT over \mathbf{Z}_2 is simply the Hadamard gate H . The QFT over \mathbf{Z}_2^m is the m -fold tensor

product of H with itself and a partial QFT acts as H on some subset of the qubits and as the identity elsewhere. An automorphism of \mathbf{Z}_2^m is a linear map $x \rightarrow Ax$ where A is an invertible $m \times m$ matrix over \mathbf{Z}_2 . An example is the CNOT operation between any two bits. An example of a quadratic function for $m = 1$ is the function $x \rightarrow i^x$ where $x \in \mathbf{Z}_2$. To verify that (4) holds, one uses that $i^{x+y} = i^x i^y (-1)^{xy}$ for every $x, y \in \mathbf{Z}_2$, where $x + y$ denotes the sum modulo 2. For $m = 2$ the function $(x, y) \rightarrow (-1)^{xy}$ is quadratic as well, as can be easily verified. The quadratic phase gates corresponding to the above functions are the single-qubit $\frac{\pi}{2}$ -phase gate and the two-qubit controlled-Z gate, respectively:

$$D = \text{diag}(1, i); \quad CZ = \text{diag}(1, 1, 1, -1). \quad (5)$$

For arbitrary m , the gate D acting on any single qubit and CZ acting on any pair of qubits are quadratic phase gates as well. We can thus conclude that every quantum circuit composed of CNOT, H , CZ and D gates is a normalizer circuit over \mathbf{Z}_2^m . We thus recover the well known class of *Clifford circuits*.

3.2 \mathbf{Z}_{2^m} and “the” QFT

Consider the group \mathbf{Z}_{2^n} consisting of integers modulo 2^n . The space \mathcal{H} is 2^n -dimensional and the standard basis is labeled by the integers from 0 to $2^n - 1$. The difference with \mathbf{Z}_2^m is that the structure of \mathbf{Z}_{2^m} does not naturally induce a factorization of \mathcal{H} into m single-qubit systems, since the decomposition (1) only contains a single cyclic group. As a consequence we will see that normalizer gates over \mathbf{Z}_{2^m} will act *globally* on \mathcal{H} , in contrast with \mathbf{Z}_2^m where we found that certain normalizer gates are single- and two-qubit gates.

The QFT over \mathbf{Z}_{2^m} is the commonly known transform occurring in e.g. Shor’s factoring algorithm. An explicit formula is obtained by replacing d_i by 2^m in (2). There are no partial QFTs since the decomposition (1) contains a single cyclic group. Consider an arbitrary $a \in \mathbf{Z}_{2^m}$ which is coprime to 2^m . Then the multiplication map $x \rightarrow ax$ is an automorphism; the coprimality condition ensures that this map is invertible [15]. Hence $|x\rangle \rightarrow |ax\rangle$ defines an automorphism gate. In section 11 we show that

$$D = \sum \omega^{bx^2+cx} |x\rangle\langle x| \quad \text{and} \quad E = \sum \gamma^{bx(x+2^m)} |x\rangle\langle x|; \quad \omega = e^{2\pi i/2^m}, \quad \gamma = \omega^{1/2} \quad (6)$$

are quadratic phase gates for arbitrary $b, c \in \mathbf{Z}_{2^m}$. In the definition of E , the quantity $bx(x+2^m)$ is computed over the integers.

4 Main result

In our main result (theorem 1 below) we show that all polynomial size normalizer circuits can be efficiently simulated classically. Before stating this result we make some notions more precise.

Generally speaking, in the context of classical simulations the task is the following: given the description of a quantum circuit, its input and measurement, reproduce the output of the computation in polynomial time on a classical computer. In the standard quantum circuit paradigm comprising local gates acting on at most d qubits for some constant d , each gate can be naturally described by specifying on which d -qubit subsystem it acts and by listing all $2^d \times 2^d$ matrix elements; the description of the entire circuit is then essentially the concatenation of all descriptions of the gates in the circuit. In the context of normalizer

circuits, however, the situation is different since these gates may act *globally* on the entire \mathfrak{g} -dimensional Hilbert space (cf. the example of \mathbf{Z}_{2^n} above). Therefore, specifying such operators in terms of their $\mathfrak{g} \times \mathfrak{g}$ matrix elements is in general exponentially inefficient and a different encoding should be used. In fact it is a priori not obvious that an efficient encoding, i.e. comprising $\text{polylog}(\mathfrak{g})$ bits, exists at all for arbitrary normalizer gates. Next we show that this is the case.

First, a partial quantum Fourier transform is simply described by the set of systems \mathcal{H}_i on which it acts nontrivially.

Second we consider automorphism gates. For every i ranging from 1 to m , denote by $e^i \in G$ the group element which has $1 \in \mathbf{Z}_{d_i}$ in its i -th component and zeroes elsewhere, where 0 in slot k represents the neutral element in \mathbf{Z}_{d_k} . Remark that $g = \sum g_i e^i$ for every $g \in G$ so that the m elements e^i generate G . Therefore every automorphism α is fully specified by the $m \times m$ matrix

$$A := [\alpha(e^1) | \cdots | \alpha(e^m)]. \quad (7)$$

We call A the matrix representation of α . Note that each column contains an element of G and can thus be described by $O(\sum \log d_i) = O(\log \mathfrak{g})$ bits. Furthermore for every decomposition (1) one has $m = O(\log \mathfrak{g})$ so that A has $O(\log \mathfrak{g})$ columns [25]. Therefore, A yields an efficient description of U_α .

Third, letting ξ be an arbitrary quadratic function, in section 6.3 we will show that

- Q1.** ξ is completely determined by its action on the elements e^i and $e^i + e^j$;
- Q2.** $\xi(g)^{2\mathfrak{g}} = 1$ for every $g \in G$. Thus there exist $n(g) \in \mathbf{Z}_{2\mathfrak{g}}$ such that $\xi(g) = e^{\pi i n(g)/\mathfrak{g}}$.

It follows that the $O(m^2)$ integers $n(e^i)$ and $n(e^i + e^j)$ comprise an efficient description of ξ and thus of the associated quadratic phase gate.

Henceforth we will assume that all normalizer gates are specified in terms of the descriptions given above, which will be called their *standard encodings*. The standard encoding of a normalizer circuit consists of the concatenation of the standard encodings of the normalizer gates of which it is made up.

Finally, we will consider the following family of input states. If K is a subgroup of G and $x \in G$, the associated coset state is

$$|K + x\rangle = \frac{1}{\sqrt{K}} \sum |k + x\rangle, \quad (8)$$

where the sum is over all $k \in K$. The standard encoding of this coset state is given by a generating set of K of size $\text{polylog}(\mathfrak{g})$ (henceforth called a polynomial size generating set) together with the element x . This description is efficient. Remark that in particular every standard basis state is a coset state.

Theorem 1 (Classical simulations of normalizer circuits). *Let $G = \mathbf{Z}_{d_1} \times \cdots \times \mathbf{Z}_{d_m}$ be a finite Abelian group. Consider a polynomial size normalizer circuit over G acting on a coset input state, where both circuit and input are specified in terms of their standard encodings as described above. The circuit is followed by a measurement in the standard basis. Then there exists an efficient classical algorithm to sample the corresponding output distribution.*

Theorem 1 generalizes a well known result in the theory of classical simulations viz. the Gottesman-Knill theorem [6]. The latter regards classical simulations of Clifford circuits.

The connection between normalizer circuits and Clifford circuits is obtained by considering the group \mathbf{Z}_2^m as in section 3.1. Theorem 1 thus shows that every polynomial-size Clifford circuit acting on a coset input state and followed by a computational basis measurement, can be simulated efficiently classically. This recovers (a variant of) the Gottesman-Knill theorem. Analogously, considering \mathbf{Z}_d^m where d is any constant, theorem 1 recovers previously known generalizations of the Gottesman-Knill theorem from qubits to d -level systems [18, 19].

The novel content of theorem 1 regards all groups which *cannot* be decomposed as a direct product \mathbf{Z}_d^m with d constant. An interesting case is the single cyclic group \mathbf{Z}_{2^m} as in section 3.2. Theorem 1 implies that every quantum circuit composed of $\text{poly}(m)$ QFTs over \mathbf{Z}_{2^m} , multiplication gates $|x\rangle \rightarrow |ax\rangle$ and diagonal gates D and E , when acting on a coset input state and followed by standard basis measurement, can be simulated classically in $\text{poly}(m)$ time. More generally, one may also consider e.g. an n -fold direct product $\mathbf{Z}_{2^n}^m$ and construct the corresponding normalizer circuits. In this case “efficient” is synonymous to “in $\text{poly}(n, m)$ time”. Theorem 1 implies that polynomial size normalizer circuits over this group can be efficiently simulated classically. We refer to section 11 for more examples of normalizer gates.

Finally we point out two additional features of our results. First, theorem 1 will yield an efficient algorithm to sample the output distribution Π of any normalizer circuit with *perfect* accuracy. Second, we will fully characterize Π as follows: *There exists $x \in G$ and a subgroup $H \subseteq G$ such that Π is the uniform distribution over the coset $H + x$.* Furthermore x and a generating set of H can be computed efficiently.

5 Relation to Shor’s factoring algorithm and the HSP

We revisit Shor’s factoring algorithm [3] in light of theorem 1. Recall that this algorithm employs two quantum registers. Register 1 is a 2^m -dimensional Hilbert space with standard basis states $|x\rangle$ with $x \in \mathbf{Z}_{2^m}$. Register 2 is N -dimensional with basis states $|y\rangle$ with $y \in \mathbf{Z}_N$ where N is the number one wishes to factor. Fixing also $a \in \mathbf{Z}_N$ coprime to N , Shor’s algorithm consists of the following steps [3, 2, 16].

- S1.** Initialize the total system in the state $|0, 0\rangle$.
- S2.** Apply the QFT over \mathbf{Z}_{2^m} to register 1.
- S3.** Apply the unitary operator U_{me} defined by $|x, y\rangle \rightarrow |x, y + a^x\rangle$.
- S4.** Apply the QFT over \mathbf{Z}_{2^m} to register 1.
- S5.** Measure register 1 in the standard basis.

Repeating the above procedure T times for some suitable T scaling polynomially with the input size $\log N$ and suitably postprocessing the collected measurement outcomes with a classical computer yields a factor of N in polynomial time, as proved by Shor. We now consider the class of normalizer circuits over $G = \mathbf{Z}_{2^m} \times \mathbf{Z}_N$ and ask whether S1-S5 are accounted for by theorem 1. Except modular exponentiation S3, all steps are obviously covered. In section 10 we prove that U_{me} cannot be well-approximated by any normalizer circuit:

Theorem 2. *Let $N = pq$ where $p \geq 2$ and $q \geq 2$ are prime with $p \neq q$, let $a \in \mathbf{Z}_N$ be coprime to N and consider $m \geq 2$. Consider the unitary operator U_{me} defined in S3. There does not exist any (even exponential-size) normalizer circuit \mathcal{C} satisfying $\|\mathcal{C} - U_{\text{me}}\| < 1 - 2^{-1/2}$ where $\|\cdot\|$ denotes the operator norm.*

Therefore theorem 1 does not lead to an efficient classical simulation of Shor’s algorithm. But what is now “the” crucial element accounting for the speed-up of this algorithm? Is it the QFT—after all the only other component in the algorithm is modular exponentiation, which is a “classical” operation as it can entirely realized by e.g. a circuit of Toffoli gates? Or is it modular exponentiation—after all this is the only component not covered by the formalism of normalizer circuits? The point is that such questions do not have absolute answers as they are only meaningful in a relative setting: relative to the class of Toffoli circuits, the QFT is the essential ingredient that gives the factoring algorithm its power. By the same token, relative to the class of normalizer circuits, the essential ingredient is precisely modular exponentiation. The power of quantum factoring lies precisely in its ingenious combination of both elements.

We provide further connections between normalizer circuits and quantum factoring. Broadening our scope, we treat the hidden subgroup problem (HSP) for arbitrary Abelian groups [4, 2, 16]. In the HSP one has oracle access to a function $f : G \rightarrow \mathbf{Z}_N$ for some Abelian group G and for some N , where it is promised that there exists an (unknown) subgroup $H \subseteq G$ such that $f(g) = f(h)$ iff $g - h \in H$. The goal is to compute a generating set of H . Whereas the best classical algorithm requires $O(\mathfrak{g})$ operations, there exists a quantum algorithm which solves the problem in $\text{polylog}(\mathfrak{g})$ time. Very similar to Shor’s algorithm, the HSP algorithm uses two registers. Register 1 is \mathfrak{g} -dimensional with standard basis states $|g\rangle$ where $g \in G$, and register 2 is N -dimensional with basis states $|y\rangle$ with $y \in \mathbf{Z}_N$. Letting F denote the QFT over G , the steps in the HSP algorithm are fully analogous to S1-S5 as follows: initialize in $|0, 0\rangle$; apply F to register 1; apply $U_f : |g, y\rangle \rightarrow |g, y + f(g)\rangle$; apply F to register 1; measure register 1.

Similar to Shor’s algorithm, the quantum circuit for the HSP is not realized by any normalizer circuit because of U_f . However, let us consider the state ρ of register 1 after application of U_f . It is well known (and easily verified) that ρ is the uniform mixture of all coset states of H :

$$\rho = \frac{1}{\mathfrak{g}} \sum_{x \in G} |H + x\rangle \langle H + x|. \quad (9)$$

Now remark that the only remaining operation in the HSP algorithm is the QFT, which is a normalizer gate. Revisiting theorem 1 we find that a computation with input ρ followed by the application of the QFT can be simulated efficiently classically (by generating a random $x \in G$ and by classically simulating the QFT acting on the input $|H + x\rangle$). This appears to imply that the quantum algorithm for the HSP can be simulated efficiently classically by virtue of theorem 1. However this conclusion is incorrect: the crucial point is that in the HSP the state ρ is *unknown* as H is unknown—indeed the entire purpose of the HSP is to identify H ! Therefore theorem 1 cannot be applied in this setting.

In short, the HSP can be embedded in the *same* computational model considered in theorem 1, which has normalizer circuits acting on (mixtures of) coset state(s) as its main ingredients. The crucial distinction is that in the HSP the input state ρ is unknown and the goal is to identify it, as in a tomography problem. This viewpoint also highlights the important role of the *first* stage of the HSP algorithm whose purpose is the preparation of the unknown ρ via application of U_f – even though this first stage appears to be a mere “classical” step! Another implication is that we cannot expect to learn much about the speed-up achieved by the HSP algorithm (and thus in particular Shor’s algorithm) by quantifying, say, how much entanglement is generated during the subroutine $\rho \rightarrow F\rho F^\dagger$.

Indeed the exact same process with a fully known but otherwise identical input state ρ can be simulated classically.

6 Preliminaries on Abelian groups

Here we discuss concepts in group theory such as characters, endomorphisms and quadratic functions, as well as the computational complexity of some group theoretic tasks.

6.1 Characters and orthogonal groups

We review some basic representation theory of finite Abelian groups; see [17] for more details. The set \mathcal{F} of all complex functions $f : G \rightarrow \mathbf{C}$ is well known to be a \mathfrak{g} -dimensional Hilbert space where the inner product between two functions f_1 and f_2 is defined by $\langle f_1, f_2 \rangle := \sum \bar{f}_1(g) f_2(g)$; here the sum is over all $g \in G$ and $\bar{f}_1(g)$ is the complex conjugate of $f_1(g)$. A character of G is a function $\chi \in \mathcal{F}$ satisfying $\chi(g) \neq 0$ for every g and

$$\chi(g+h) = \chi(g)\chi(h) \quad \text{for every } g, h \in G. \quad (10)$$

Equivalently, a character is a homomorphism from G to the multiplicative group of nonzero complex numbers. The trivial character maps $g \rightarrow 1$ for every g . For every $g = g_1 \cdots g_m$, define the following complex function on G :

$$\chi_g : h = h_1 \cdots h_m \rightarrow \exp 2\pi i \left(\frac{g_1 h_1}{d_1} + \cdots + \frac{g_m h_m}{d_m} \right). \quad (11)$$

Lemma 1 (Characters). *For all $g, h \in G$ and for all characters χ and ψ the following holds:*

- (a) *If $\chi \neq \psi$ then $\langle \chi, \psi \rangle = 0$ i.e. these characters are orthogonal.*
- (b) *χ_g is a character of G .*
- (c) *$\chi_g(h) = \chi_h(g)$.*
- (d) *$\chi_{g+h} = \chi_g \chi_h$.*
- (e) *Define $S(g) := \sum_{h \in G} \chi_g(h)$. Then $S(0) = \mathfrak{g}$ and $S(g) = 0$ for every $g \neq 0$.*
- (f) *Every character is a quadratic function.*
- (g) *The QFT over G satisfies $F = \frac{1}{\sqrt{\mathfrak{g}}} \sum \chi_g(h) |g\rangle \langle h|$, where the sum is over all $g, h \in G$.*
- (h) *Every character of G is one of the functions χ_g i.e. G has precisely \mathfrak{g} characters.*

Proof: since χ is a homomorphism one has $\chi(0) = 1$. Since $\mathfrak{g}g = 0$ for every $g \in G$ and since χ is a character it follows that $1 = \chi(0) = \chi(\mathfrak{g}g) = \chi(g)^\mathfrak{g}$. This shows that $\chi(g)$ is a complex phase for every $g \in G$. If χ is different from ψ there exists $g_0 \in G$ such that $\chi(g_0) \neq \psi(g_0)$ and thus, since $\chi(g_0)$ is a complex phase, $\bar{\chi}(g_0)\psi(g_0) \neq 1$. Then

$$\begin{aligned} \langle \chi, \psi \rangle &= \sum \bar{\chi}(g) \psi(g) = \sum \bar{\chi}(g + g_0) \psi(g + g_0) \\ &= \bar{\chi}(g_0) \psi(g_0) \sum \bar{\chi}(g) \psi(g) = \bar{\chi}(g_0) \psi(g_0) \langle \chi, \psi \rangle \end{aligned} \quad (12)$$

where sums are over all $g \in G$, and claim (a) follows. Claims (b)-(c)-(d) follow straightforwardly from definition (11). To prove (e) note that $S(g)$ coincides with the inner product between the character χ_g and the trivial character. Furthermore χ_g is the trivial character iff $g = 0$. Using (a) then yields (e). To prove (f) remark that (4) is satisfied by taking B to

be the trivial bilinear function which sends all (g, h) to 1. Claim (g) is obtained straightforwardly by using the definitions of χ_g and F . Let \mathcal{F} denote the \mathfrak{g} -dimensional space of all complex functions on G as above. Since every two characters are orthogonal owing to (a), there can be at most \mathfrak{g} characters. Owing to (b) all functions χ_g are characters; furthermore $\chi_g = \chi_h$ iff $g = h$ so that there are precisely \mathfrak{g} distinct functions χ_g . It follows that there can be no other characters; this proves (h). \square

For every subgroup $H \subseteq G$ define the following subset of G :

$$H^\perp := \{g \in G : \chi_g(h) = 1 \text{ for every } h \in H\}. \quad (13)$$

Lemma 2 (Orthogonal groups). *For every subgroup $H \subseteq G$ the following holds.*

- (a) H^\perp is a subgroup of G , called the orthogonal group of H .
- (b) $|H^\perp| \cdot |H| = \mathfrak{g}$.
- (c) $(H^\perp)^\perp = H$.

Proof: (a) is proved straightforwardly by using that each χ_g is a character. To prove (b), for every $g \in G$ let $\psi_g : H \rightarrow \mathbf{C}$ be the restriction of χ_g to H i.e. $\psi_g(h) = \chi_g(h)$ for every $h \in H$. The function ψ_g is easily seen to be a character of H . Define

$$\tau(g) := \sum_{h \in H} \chi_g(h) = \sum_{h \in H} \psi_g(h). \quad (14)$$

Remark that $\tau(g)$ is the inner product between the trivial character of H and the character ψ_g . Lemma 1a implies that $\tau(g) = |H|$ if ψ_g is the trivial character on H and $\tau(g) = 0$ otherwise. Furthermore the definition of H^\perp implies that ψ_g is the trivial character precisely when $g \in H^\perp$. Thus $\tau(g) = |H|$ iff $g \in H^\perp$. Now consider the state $|H\rangle$ as in (8). Then (recalling lemma 1g)

$$F|H\rangle = \frac{1}{\sqrt{g|H|}} \sum_{g \in G} \sum_{h \in H} \chi_g(h) |g\rangle = \frac{1}{\sqrt{g|H|}} \sum_{g \in G} \tau(g) |g\rangle = \frac{\sqrt{|H|}}{\sqrt{g}} \sum_{g \in H^\perp} |g\rangle. \quad (15)$$

Since F is unitary, the norm of $F|H\rangle$ is 1. Together with (15) this yields (b). Finally we prove (c). By definition one has $g \in H^{\perp\perp}$ iff $\chi_g(k) = 1$ for all $k \in H^\perp$. Consider $h \in H$. Then by definition of H^\perp one has $\chi_k(h) = 1$ for all $k \in H^\perp$. Using lemma 1c it follows that $h \in H^{\perp\perp}$ and we conclude that $H \subseteq H^{\perp\perp}$. Applying claim (b) to H and H^\perp yields $|H||H^\perp| = \mathfrak{g} = |H^\perp||H^{\perp\perp}|$ so that $|H| = |H^{\perp\perp}|$. But then we must have $H = H^{\perp\perp}$. \square

6.2 Endomorphisms

An endomorphism ϖ of G is a homomorphism from G to itself i.e. $\varpi(g+h) = \varpi(g) + \varpi(h)$ for all $g, h \in G$. Every endomorphism ϖ has an $m \times m$ matrix representation similar to (7). We characterize which matrices can occur a matrix representations of endomorphisms.

Lemma 3 (Matrix representation). *Let A be an $m \times m$ matrix where the entries in the k -th row of A belong to \mathbf{Z}_{d_k} for every k . Let $a^i \in G$ denote the i -th column of A . Then there exists an endomorphism of G with matrix representation A if and only if $d_i a^i = 0$ for every i .*

Proof: let A be the matrix representation of an endomorphism ϖ . Since $d_i e^i = 0$, it follows that $0 = \varpi(0) = \varpi(d_i e^i) = d_i \varpi(e^i)$ and so $d_i a^i = 0$. We prove the converse. Consider a matrix A satisfying $d_i a^i = 0$ for every i . Define $\varpi(g) := \sum_i g_i a^i$ for every $g = g_1 \cdots g_m \in G$. We prove that ϖ is an endomorphism. Consider arbitrary $g, h \in G$. We introduce the following notation: the sum of g_i and h_i over \mathbf{Z}_{d_i} is denoted by $g_i \oplus h_i$. The sum of these two numbers regarded as *integers* is denoted by $g_i + h_i$. With this notation, applying the definition of ϖ one obtains

$$\varpi(g + h) = \sum (g_i \oplus h_i) a^i. \quad (16)$$

$$\varpi(g) + \varpi(h) = \sum g_j a^j + \sum h_i a^i = \sum (g_i + h_i) a^i. \quad (17)$$

Since the i -th component of $g + h$ is $g_i \oplus h_i = g_i + h_i \bmod d_i$, there exists an integer q_i such that $g_i \oplus h_i = g_i + h_i + q_i d_i$. Since $d_i a^i = 0$, it follows that $(g_i + h_i) a^i = (g_i \oplus h_i) a^i$. Combining this last equation with (16) and (17) shows that ϖ is an endomorphism. \square

Consider an endomorphism ϖ , fix $g \in G$ arbitrarily and consider the map $\varphi_g = \chi_g \cdot \varpi$:

$$\varphi_g : x \in G \rightarrow \chi_g(\varpi(x)). \quad (18)$$

Using that χ_g is a character and that ϖ is an endomorphism one shows that φ_g is also a character. Recalling lemma 1h there must hence exist a unique \hat{g} such that $\varphi_g = \chi_{\hat{g}}$. Let $\hat{\varpi}$ denote the map which sends g to \hat{g} . Using lemma 1d it is straightforward to show that $\hat{\varpi}$ is also an endomorphism. Indeed for every $x \in G$ one has

$$\varphi_{g+h}(x) = \chi_{g+h}(\varpi(x)) = \chi_g(\varpi(x)) \chi_h(\varpi(x)) = \varphi_g(x) \varphi_h(x) \quad (19)$$

implying that $\varphi_{g+h} = \varphi_g \varphi_h$ and thus $\hat{\varpi}(g + h) = \hat{\varpi}(g) + \hat{\varpi}(h)$ for every $g, h \in G$. The map $\hat{\varpi}$ is henceforth called here the dual endomorphism of ϖ .

Lemma 4 (Dual matrix representation). *Let A be the matrix representation of an endomorphism ϖ . Then the matrix representation B of $\hat{\varpi}$ is given by*

$$\frac{d_k}{d_l} A_{lk} \bmod d_k = B_{kl} \quad \text{for every } k, l = 1, \dots, m. \quad (20)$$

Proof: we have $\chi_{e^l}(\varpi(e^k)) = \chi_{\hat{\varpi}(e^l)}(e^k)$ for every k, l by definition of $\hat{\varpi}$. Using that A and B are the matrix representations of ϖ and $\hat{\varpi}$, respectively, this yields

$$\exp \frac{2\pi i A_{lk}}{d_l} = \exp \frac{2\pi i B_{kl}}{d_k}. \quad (21)$$

It follows that B_{kl} is equal to $d_k A_{lk} / d_l$ up to addition of an integer multiple of d_k . This implies that

$$d_k A_{lk} / d_l \bmod d_k = B_{kl} \bmod d_k. \quad (22)$$

Since B is the matrix representation of an endomorphism, B_{kl} is an integer between 0 and $d_k - 1$. This implies that $B_{kl} \bmod d_k = B_{kl}$ and the result follows. \square

To illustrate the above result, consider a group G of the form \mathbf{Z}_d^m i.e. all cyclic groups in the decomposition of G have the same order. Then lemma 4 implies that the matrix representation of the dual endomorphism is simply the transposed matrix A^T . Finally, we remark that lemma 4 can be used to show that $(\hat{\varpi})^\wedge = \varpi$. The argument is straightforward.

6.3 Quadratic functions

We prove claims Q1-Q2 made in section 4 by means of the following lemma.

Lemma 5 (Quadratic and bilinear functions). *(a) A function B from $G \times G$ to the nonzero complex numbers is bilinear if and only if there exists an endomorphism ϖ of G such that $B(g, h) = \chi_{\varpi(g)}(h)$. Furthermore for every bilinear function B and quadratic function ξ as in (4), for every $g, h \in G$ and for every positive integer n , the following holds:*

- (b) $B(g, 0) = 1 = B(0, g)$.
- (c) $B(g, h)^{\mathfrak{g}} = 1$.
- (d) $\xi(0) = 1$.
- (e) $\xi(ng) = \xi(g)^n B(g, g)^{f(n)}$ where $f(n) = n(n-1)/2$.
- (f) $\xi(g)^{2\mathfrak{g}} = 1$.

Proof: consider an endomorphism ϖ . Using lemma 1b,d and the fact that ϖ is an endomorphism, one shows that $B(g, h) = \chi_{\varpi(g)}(h)$ is bilinear. Conversely, consider an arbitrary bilinear function B . By definition of bilinearity, for every fixed g the function $\beta_g : h \rightarrow B(g, h)$ is a character. Lemma 1h implies that there exists a unique g^* such that $\beta_g = \chi_{g^*}$. Let ϖ denote the map which sends g to g^* . Using that $g \rightarrow B(g, h)$ is also a character and lemma 1d it follows that ϖ must be an endomorphism. This proves (a). Claim (b) follows from (a) and the fact that $\varpi(0) = 0$ for every endomorphism. Claim (c) is proved by using (a) and by noting that $\chi_x(y)^{\mathfrak{g}} = 1$ for every $x, y \in G$ where the latter follows directly from definition (11). Using that $B(0, 0) = 1$ one finds

$$\xi(0) = \xi(0 + 0) = \xi(0)\xi(0)B(0, 0) = \xi(0)^2. \quad (23)$$

Recalling in addition that by definition $\xi(g) \neq 0$ for every g yields (d). Claim (e) is proved by induction on n ; the statement is obviously correct for $n = 0, 1$. Assuming (e) holds for some integer n , one has

$$\xi((n+1)g) = \xi(ng)\xi(g)B(ng, g) = \xi(ng)\xi(g)B(g, g)^n = \xi(g)^{n+1} B(g, g)^{f(n)+n} \quad (24)$$

In the first identity we used that ξ is quadratic, in the second that B is bilinear; in the final identity we used the induction hypothesis. Noting that $f(n) + n = f(n+1)$ yields the proof. Using (e) with $n = 2\mathfrak{g}$ and recalling (c) yields $\xi(2\mathfrak{g}g) = \xi(g)^{2\mathfrak{g}}$. Noting that $2\mathfrak{g}g = 0$ and recalling (d) proves (f). \square

Lemma 5f proves that $\xi(g)$ is a complex phase. Repeatedly applying the definitions of quadratic and bilinear functions yields, for every $g = \sum g_i e^i \in G$:

$$\xi(g) = \prod_i \xi(g_i e^i) \prod_{i < j} B(e^i, e^j)^{g_i g_j} \quad (25)$$

$$B(e^i, e^j) = \xi(e^i + e^j) \overline{\xi(e^i) \xi(e^j)}. \quad (26)$$

Furthermore, lemma 5e implies

$$\xi(g_i e^i) = \xi(e^i)^{g_i} B(e^i, e^i)^{f(g_i)}. \quad (27)$$

We conclude that $\xi(g)$ is fully determined by the values $\xi(e^i)$ and $\xi(e^i + e^j)$.

If ξ is a quadratic function where the associated B is given as in lemma 5a, we will simply say that ξ is a quadratic function with endomorphism ϖ . For every quadratic function, the

associated B and ϖ are unique. Uniqueness of B follows from (4). To show uniqueness of ϖ , assume that $B(g, h) = \chi_{\varphi(g)}(h)$ for some other endomorphism φ . Then

$$1 = B(0, h) = B(g - g, h) = B(g, h)B(-g, h) = \chi_{\varpi(g)}(h)\chi_{-\varphi(g)}(h) = \chi_{\varpi(g)-\varphi(g)}(h), \quad (28)$$

for all $g, h \in G$. This implies that $\varpi = \varphi$.

6.4 Computations in Abelian groups

We show that several problems involving the concepts introduced earlier in this section have efficient classical algorithms. In the following the *size* of an integer is the number of bits in its binary expansion. Recall that for every group (1) one has $m = O(\log \mathfrak{g})$. Remark also that the size of each d_i is $O(\log \mathfrak{g})$ since $d_i \leq \mathfrak{g}$ and thus for every $g = g_1 \cdots g_m \in G$ the size of each g_i is $O(\log \mathfrak{g})$ as well.

We start by discussing some elementary tasks. Given two integers a and b of size at most l , their sum, product as well as the remainder $a \bmod b$ can be computed in $\text{poly}(l)$ time. Hence the sum $g + h$ can be computed in $\text{polylog}(\mathfrak{g})$ time for every $g, h \in G$ by computing the m remainders $g_i + h_i \bmod d_i$. Similarly, ng can be computed in $\text{polylog}(\mathfrak{g}, n)$ time for every integer n by computing the remainders $ng_i \bmod d_i$.

Let ϖ be an endomorphism given in terms of its matrix representation. Remarking that $\varpi(g) = \sum g_i \varpi(e^i)$ for every $g = g_1 \cdots g_m$, the group element $\varpi(g)$ can be computed efficiently on input of g . Lemma 4 shows that the matrix representation of $\hat{\varpi}$ can be computed efficiently as well.

Let ξ be a quadratic function given in terms of its standard encoding. Consider the associated B as in (4) and ϖ as in lemma 5a. Then the matrix representation A of ϖ can be computed efficiently: one first computes the quantities $B(e^k, e^l)$ using (26). Using

$$B(e^k, e^l) = \chi_{\varpi(e^k)}(e^l) = e^{2\pi i A_{lk}/d_l} \quad (29)$$

yields the matrix element A_{lk} . Furthermore the function value $\xi(g)$ can be computed efficiently on input of g : first compute all quantities $B(e^i, e^j)$ using (26), then compute $\xi(g_i e^i)$ using (27), finally compute $\xi(g)$ using (25).

Next we discuss more involved tasks:

Theorem 3 (Efficient computations in Abelian groups). *Consider a finite Abelian group (1). Then the following problems can be solved in $\text{polylog}(\mathfrak{g})$ time:*

- (a) *Given a generating set of a subgroup H , compute a generating set of H^\perp .*
- (b) *Given $h^1, \dots, h^r \in G$ with $r = \text{polylog}(\mathfrak{g})$ and $s_1, \dots, s_r \in \mathbf{Z}_{\mathfrak{g}}$, determine whether there exists $g \in G$ satisfying*

$$\chi_{h^k}(g) = e^{2\pi i s_k/\mathfrak{g}} \quad \text{for all } k \quad (30)$$

and, if yes, compute such a solution g .

- (c) *Given the matrix representation of an automorphism α , compute the matrix representation of α^{-1} .*

Property (a) is known in the quantum algorithms community, cf. the Abelian HSP algorithm [5]. Property (b) is similar to (a). It is however not clear to the author if (c) is known. A proof of theorem 3 is given in appendix B. For now we mention that a central ingredient in the proof is the fact that systems of linear diophantine equations can be solved efficiently, as recalled next.

Theorem 4 (Linear Diophantine equations [20, 21, 22, 23]). *Let A be an $n \times m$ matrix and let b be an n -dimensional vector, both with integer entries of size at most k . Both A and b are specified by listing all of their entries in binary. Then there exists an algorithm with running time $\text{poly}(n, m, k)$ which decides whether the equation $Ax = b$ has an integer solution x . Furthermore, if a solution exist, then there exist integer vectors x^0 and x^1, \dots, x^r , for some $r \leq m$, such that:*

- (a) $Ax^0 = b$.
- (b) $Ax^i = 0$ for all $i = 1, \dots, r$.
- (c) *Every integer solution to $Ax = b$ has the form $x = x^0 + y$ where y is an integer linear combination of the x^i .*

Furthermore all components of x^0 and the x^i have size $\text{poly}(n, m, k)$ and can be computed in $\text{poly}(n, m, k)$ time.

7 Pauli and Clifford operators over Abelian groups

Here we consider the Hilbert space \mathcal{H} associated with G as in section 2 and define several families of operators acting on this space.

For every $g \in G$ define the pair of operators

$$X(g) = \sum |h + g\rangle\langle h|; \quad Z(g) = \sum \chi_g(h) |h\rangle\langle h| \quad (31)$$

where the sums are over all $h \in G$. Remark that $X(g)$ is a permutation matrix and that $Z(g)$ is a diagonal unitary operator. Let $a \in \mathbf{Z}_{2\mathfrak{g}}$ and consider the following $2\mathfrak{g}$ -th root of unity $\gamma := e^{i\pi/\mathfrak{g}}$. An operator of the form $\sigma = \gamma^a Z(g)X(h)$ will be called a Pauli operator over G , or Pauli operator in short, as this definition generalizes the standard notion of Pauli matrices defined on qubit systems [26]. The triple (a, g, h) is called the label of σ . It is important that the label constitutes a description of σ comprising $O(\log \mathfrak{g})$ bits i.e. this description is efficient. The following commutation relations are verified straightforwardly by applying (31):

$$\begin{aligned} X(g)X(h) &= X(g+h) = X(h)X(g), \\ Z(g)Z(h) &= Z(g+h) = Z(h)Z(g), \\ Z(g)X(h) &= \chi_g(h)X(h)Z(g). \end{aligned} \quad (32)$$

The first equation in (32) implies that the set $\{X(g) : g \in G\}$ is a group w.r.t. the standard operator product. In fact, this group is isomorphic to G : the isomorphism is the map $g \rightarrow X(g)$. Similarly, $\{Z(g) : g \in G\}$ is a group isomorphic to G .

Elementary manipulations with Pauli operators are computationally efficient in the following sense:

Lemma 6 (Products and powers of Pauli operators). *Consider Pauli operators σ and τ and a positive integer n . Then $\sigma\tau$, σ^n and σ^\dagger are also Pauli operators, the labels of which can be computed in $\text{polylog}(\mathfrak{g}, n)$ time on input of n and the labels of σ and τ . Moreover, $\sigma^\dagger = \sigma^{2\mathfrak{g}-1}$.*

Proof: let (a, g, h) and (b, x, y) denote the labels of σ and τ , respectively. Using the commutation relations (32) one finds

$$\sigma\tau = \gamma^{a+b} \cdot \chi_h(x) \cdot X(g+x)Z(h+y). \quad (33)$$

The overall phase on the r.h.s. of (33) has the form γ^s for some $s \in \mathbf{Z}_{2\mathfrak{g}}$, so that $\sigma\tau$ is indeed a Pauli operator. It is straightforward that s as well as $g+x$ and $h+y$ can be computed efficiently. The other claims in the lemma are proved analogously. \square

Lemma 6 implies that the set of all Pauli operators forms a finite group, called here the Pauli group (over G). A unitary operator U on \mathcal{H} is called a Clifford operator (over G) if U maps the Pauli group onto itself under the conjugation map $\sigma \rightarrow U\sigma U^\dagger$. It is easy to see that the set of all Clifford operators forms a group. Formally speaking, the Clifford group is the normalizer of the Pauli group in the full unitary group acting on \mathcal{H} . We now show that normalizer gates and Clifford operators are fundamentally interrelated.

Theorem 5 (Normalizer gates are Clifford). *Every normalizer gate is a Clifford operator. Furthermore let U be a normalizer gate specified in terms of its standard classical encoding as in section 2 and let σ be a Pauli operator specified in terms of its label. Then the label of $U\sigma U^\dagger$ can be computed in $\text{polylog}(\mathfrak{g})$ time.*

Proof: we show that $UX(g)U^\dagger$ and $UZ(g)U^\dagger$ are Pauli operators with efficiently computable labels for every g . The proof of the theorem then follows from the fact that every Pauli operator is given as a product $\gamma^a X(g)Z(h)$ and by using lemma 6.

Let F be the QFT over G , consider an automorphism α and a quadratic function ξ with endomorphism ϖ . We describe the action of F , U_α and D_ξ on $X(g)$ and $Z(g)$ under conjugation; α^{-*} is shorthand for $(\alpha^{-1})^*$:

$$\begin{aligned} F &: X(g) \rightarrow Z(g); & Z(g) &\rightarrow X(-g) \\ U_\alpha &: X(g) \rightarrow X(\alpha(g)); & Z(g) &\rightarrow Z(\alpha^{-*}(g)) \\ D_\xi &: X(g) \rightarrow \xi(g)X(g)Z(\varpi(g)); & Z(g) &\rightarrow Z(g) \end{aligned} \quad (34)$$

We prove that these are the correct actions. One has

$$FX(g)F^\dagger|h\rangle = \frac{1}{\mathfrak{g}} \sum_{k,l} \overline{\chi_k(h)} \chi_{k+g}(l) |l\rangle = \frac{1}{\mathfrak{g}} \sum_l \chi_g(l) \sum_k \chi_k(l-h) |l\rangle \quad (35)$$

$$= \frac{1}{\mathfrak{g}} \sum_l \chi_g(l) \sum_k \chi_{l-h}(k) |l\rangle = \chi_g(h) |h\rangle = Z(g)|h\rangle. \quad (36)$$

The first identity is obtained using lemma 1g; using $\bar{\chi}_k(h) = \chi_k(-h)$ and lemma 1b, d yields the second identity; the third identity uses lemma 1c; the fourth is proved using lemma 1e. The action of the Fourier transform on $Z(g)$ is computed analogously and the argument is omitted.

Using that $U_\alpha^\dagger = U_{\alpha^{-1}}$ yields

$$U_\alpha X(g) U_\alpha^\dagger |h\rangle = |\alpha(\alpha^{-1}(h) + g)\rangle = |h + \alpha(g)\rangle = X(\alpha(g)) |h\rangle. \quad (37)$$

The action on $Z(g)$ is computed analogously by applying the definition of the dual automorphism. Finally, one has

$$U_\xi X(g) U_\xi^\dagger |h\rangle = \xi(g+h) \overline{\xi(h)} |g+h\rangle = \xi(g) B(g,h) |g+h\rangle \quad (38)$$

$$= \xi(g) \chi_{\varpi(g)}(h) |g+h\rangle = \xi(g) X(g) Z(\varpi(g)) |h\rangle, \quad (39)$$

where we in the second identity we have used (4) and in the third lemma 5a.

We now show that the actions of normalizer gates on Pauli operators can be computed efficiently. The argument heavily relies on theorem 3 as well as the discussion above it. The action of F can obviously be computed efficiently. As for U_α , recall that the action $g \rightarrow \alpha(g)$ can be computed efficiently. As the matrix representations of inverse and dual automorphisms can be computed efficiently, the action $g \rightarrow \alpha^{-*}(g)$ can be computed efficiently as well. As for D_ξ , recall that both $\xi(g)$ and $\varpi(g)$ can be computed efficiently on input of g together with the standard representation of ξ .

Finally, analogous to the arguments above, it can be shown that the action of partial QFTs can be computed efficiently as well; this argument is omitted. \square

Variants of the above result were previously known for groups of the form \mathbf{Z}_d^m with d constant (see [18, 19] and references within), although those works do not use the terminology of normalizer gates, automorphism gates, quadratic phase gates, etc.

We conclude by remarking that every Pauli operator can be realized as a polynomial size (in fact: constant size) normalizer circuit, up to a global phase. To see this, first note that $Z(g)$ is a quadratic phase gate since the functions χ_g are quadratic functions (recall lemma 1f). As for $X(g)$, recall from (34) that $X(g) = F^\dagger Z(g) F$. But then every Pauli operator is realized as a constant-size normalizer circuit as well.

8 M-states

Next we recall results of [24] where a formalism was introduced to describe and classically simulate a class of quantum states called M-states. The brief account given here is self-contained and will suffice for our purposes.

As before we let G be an Abelian group with associated Hilbert space \mathcal{H} . A unitary operator U on \mathcal{H} is *monomial* if $U = PD$ for some diagonal matrix D and some permutation matrix P , both relative to the standard basis. A state $|\psi\rangle$ is an *M-state* if there exists a group \mathcal{G} consisting entirely of monomial unitary matrices such that $|\psi\rangle$ is the unique eigenvector with eigenvalue 1 of all elements of \mathcal{G} simultaneously. The group \mathcal{G} is said to be a monomial stabilizer group of $|\psi\rangle$. In theorem 6 below we characterize the structure of an M-state $|\psi\rangle$ in terms of the properties of \mathcal{G} . To do so we first develop some further concepts. Consider $U \in G$ and $g \in G$; then, since U is monomial, there exists $h \in G$ such that $U|g\rangle \propto |h\rangle$. The group thus \mathcal{G} naturally acts on the elements of G via the map $g \rightarrow h$. The orbit of g under this action is the set of all h that can be reached by applying all possible elements from \mathcal{G} :

$$O_g = \{h : \exists U \in \mathcal{G} \text{ s.t. } U|g\rangle \propto |h\rangle\}. \quad (40)$$

Finally, for every $g \in G$ let \mathcal{G}_g be the subset of \mathcal{G} consisting of all $U \in \mathcal{G}$ satisfying $U|g\rangle \propto |g\rangle$ i.e. U acts trivially on g ; this subset is easily seen to be a subgroup of \mathcal{G} .

Theorem 6 (M-states [24]). *Let $|\psi\rangle$ be an M-state with monomial stabilizer group \mathcal{G} . Then the following holds:*

- (a) $|\psi\rangle$ is a uniform superposition. This means that all non-zero amplitudes $\langle g|\psi\rangle$ are equal in modulus.
- (b) Let $S \subseteq G$ be the support of $|\psi\rangle$ i.e. S is the collection of all g with $\langle g|\psi\rangle \neq 0$. Then S coincides with one of the orbits O_x .
- (c) Consider an arbitrary $x \in G$ and let V_1, \dots, V_r be generators of \mathcal{G}_x . Then one has

$$S = O_x \quad \text{iff} \quad V_i|x\rangle = |x\rangle \text{ for all } i. \quad (41)$$

For completeness, theorem 6 is proved in appendix C. The relevance of this result for the proof of theorem 1 is the following. We will show that the output state $|\psi\rangle$ of any normalizer circuit is always an M-state. Owing to theorem 6a-b, a standard basis measurement on $|\psi\rangle$ yields a uniformly random element of O_x . We will thus prove theorem 1 by showing that the orbit O_x can be determined efficiently (using theorem 6c) and, subsequently, that a random element in O_x can be generated classically efficiently as well.

9 Proof of theorem 1

Consider a polynomial size normalizer circuit \mathcal{C} acting on the input $|K + x\rangle$. Note that $|K + x\rangle = X(x)|K\rangle$. Since $X(x)$ can be realized as a constant-size normalizer circuit (recall section 7), this operator may be absorbed into \mathcal{C} . Thus henceforth in this section we consider inputs of the form $|K\rangle$, without loss of generality.

9.1 The input state

We start by characterizing the input state as a joint eigenvector of a family of Pauli operators. We say that a state $|\psi\rangle$ is a “+1 common eigenvector” of a set \mathcal{S} of operators if the state is an eigenvector with eigenvalue 1 for every operator in the set; $|\psi\rangle$ is the “unique +1 common eigenvector” of \mathcal{S} if it is, up to a global phase, the only vector with this feature.

Lemma 7. *The set \mathcal{G}_0 consisting of all operators $Z(u)X(v)$ with $u \in K^\perp$ and $v \in K$ is an Abelian group. Furthermore $|K\rangle$ is the unique +1 common eigenvector of \mathcal{G}_0 .*

Proof: using the definitions of Pauli operators and their commutation relations it is easily verified that \mathcal{G}_0 is an Abelian group and that $|K\rangle$ is a +1 common eigenvector of \mathcal{G}_0 . To show uniqueness, let $|\psi\rangle$ be an arbitrary +1 common eigenvector of \mathcal{G}_0 , so that

$$\langle g|Z(u)X(v)|\psi\rangle = \langle g|\psi\rangle \quad \text{for every } u \in K^\perp, v \in K, g \in G. \quad (42)$$

Using (31) this implies that

$$\chi_u(g)\langle g + v|\psi\rangle = \langle g|\psi\rangle. \quad (43)$$

For every $g \notin K$ there exists $u \in K^\perp$ such that $\chi_u(g) \neq 1$: indeed, if this were not the case then g would belong to $K^{\perp\perp}$; recalling that $K^{\perp\perp} = K$ (see lemma 2c) then contradicts with the assumption that $g \notin K$. It then follows from (43) with $v = 0$ that $\langle g|\psi\rangle = 0$ for every $g \notin K$. Further, taking $u = 0 = g$ and $v \in K$ arbitrary in (43) implies that $\langle v|\psi\rangle = \langle 0|\psi\rangle$. This shows that $|\psi\rangle$ must be proportional to $|K\rangle$. \square

9.2 The output state is an M-state

Lemma 7 implies that the final state $\mathcal{C}|K\rangle =: |\psi\rangle$ is the unique +1 common eigenvector of the group \mathcal{G} obtained by applying the conjugation map $A \rightarrow \mathcal{C}A\mathcal{C}^\dagger$ to all elements of \mathcal{G}_0 . Theorem 5 shows that every element of \mathcal{G} is a Pauli operator. Remark that every Pauli operator σ is monomial and unitary as it is given as the product of a diagonal unitary matrix $\gamma^a Z(g)$ and a permutation matrix (and hence unitary) $X(h)$. Thus $|\psi\rangle$ is an M-state with monomial stabilizer group \mathcal{G} .

Theorem 6 implies that $|\psi\rangle$ is a uniform superposition and that the support of $|\psi\rangle$ is one of the orbits O_x as in (40). Next we characterize these orbits. If σ is a Pauli operator with label (a, g, h) , we call g the “ Z -component” and h the “ X -component” of σ . Denote the X -component formally by $\varphi(\sigma) := h$. Now let $H \subseteq G$ be the image of \mathcal{G} under the map φ i.e. H is the set of all X -components of the elements of \mathcal{G} . The commutation relations (32) readily yield

$$\varphi(\sigma\tau) = \varphi(\sigma) + \varphi(\tau) \quad \text{for all } \sigma, \tau \in \mathcal{G}. \quad (44)$$

This implies that H is a subgroup of G . Using that $\sigma|x\rangle \propto |x+h\rangle$ for every $\sigma \in \mathcal{G}$ with label (a, g, h) , one straightforwardly shows that

$$O_x = x + H \quad (45)$$

for every $x \in G$. Thus these orbits are precisely the cosets of the subgroup H . Theorem 6a-b now immediately implies:

Corollary 1. *There exists $x^0 \in G$ s.t. $|\psi\rangle$ is a uniform superposition with support $H + x^0$.*

Let Π denote the probability distribution resulting from measuring $|\psi\rangle$ in the standard basis. We have arrived at a crucial insight for the classical simulation we aim to achieve:

Corollary 2. *Π is the uniform distribution over the coset $H + x^0$.*

9.3 Computing a generating set of H

Next we show that a generating set of H can be computed efficiently. Let $\{u^1, \dots, u^k\}$ generate K with $k = O(\log g)$. This set is assumed to be specified as an input. Owing to theorem 3a, a generating set $\{u^{k+1}, \dots, u^n\}$ of K^\perp can be computed efficiently, for some $n = O(\log g)$. It follows that \mathcal{G}_0 is generated by the operators

$$X(u^1), \dots, X(u^k), Z(u^{k+1}), \dots, Z(u^n). \quad (46)$$

Consequently, a generating set of \mathcal{G} is obtained by conjugating these operators with \mathcal{C} :

$$\sigma_i := \begin{cases} \mathcal{C}X(u^i)\mathcal{C}^\dagger & \text{for every } i = 1, \dots, k \\ \mathcal{C}Z(u^i)\mathcal{C}^\dagger & \text{for every } i = k+1, \dots, n. \end{cases} \quad (47)$$

The σ_i mutually commute since they are obtained by simultaneously conjugating a set of mutually commuting operators. Theorem 5 shows that every σ_i is a Pauli operator with efficiently computable label, say (a_i, g^i, h^i) . Remark that (44) implies that $\varphi : \mathcal{G} \rightarrow G$ is a group homomorphism. Since \mathcal{G} is generated by the σ_i and since φ is a homomorphism, the image of φ is generated by the elements $\varphi(\sigma_i) = h^i$. But this image is precisely the subgroup H .

9.4 Computing x^0

Next we show that a suitable x^0 as in corollary 1 can be computed efficiently. Our approach will be to use theorem 6c. To this end we first characterize the subgroups $\mathcal{G}_x \subseteq \mathcal{G}$ which will turn out to have a particularly simple form; in particular we will show that all these groups are *equal*. Define the set \mathcal{D} to consist of all $\sigma \in \mathcal{G}$ that, up to some global phase, have the form $Z(g)$ for some $g \in G$. It is easily verified that this set is a subgroup of \mathcal{G} , called the diagonal subgroup.

Lemma 8. *Every group \mathcal{G}_x is equal to the diagonal subgroup \mathcal{D} .*

Proof: since every $D \in \mathcal{D}$ is a diagonal operator, one has $D|x\rangle \propto |x\rangle$, showing that $\mathcal{D} \subseteq \mathcal{G}_x$. Conversely, if $\sigma \in \mathcal{G}_x$ has label (a, g, h) then $\sigma|x\rangle \propto |x+h\rangle$. Since $\sigma \in \mathcal{G}_x$ the state $|x\rangle$ is an eigenvector of σ ; this can only be true if $h = 0$, showing that $\sigma \in \mathcal{D}$. \square

Lemma 9. *The labels of a generating set of \mathcal{D} can be computed efficiently.*

Proof: since the σ_i mutually commute, the map

$$T : k = k_1 \cdots k_n \in \mathbf{Z}^n \rightarrow \sigma_1^{k_1} \cdots \sigma_n^{k_n} \in \mathcal{G} \quad (48)$$

satisfies $T(k + k') = T(k)T(k')$. This shows that T is a group homomorphism from the additive group \mathbf{Z}^n to \mathcal{G} . Consider the inverse image of \mathcal{D} under T :

$$D = \{k \in \mathbf{Z}^n : T(k) \in \mathcal{D}\}. \quad (49)$$

Since T is a homomorphism, D is a subgroup of \mathbf{Z}^n . Furthermore we have $T(D) = \mathcal{D}$: the inclusion \subseteq holds by definition of D ; to show \supseteq , note that the σ_i generate \mathcal{G} so that every element in \mathcal{G} can be written as $T(k)$ for some k ; therefore every element in \mathcal{D} can be written as $T(k)$ with $k \in D$. Our strategy to obtain a generating set of \mathcal{D} will be to first compute a generating set of D . Since $T(D) = \mathcal{D}$, the images of these generators under T will generate \mathcal{D} .

Applying (32) one finds

$$T(k) \propto Z \left(\sum k_i g^i \right) X \left(\sum k_i h^i \right). \quad (50)$$

It follows that $k \in D$ if and only if $\sum k_i h^i = 0$. This last identity is satisfied iff

$$k_1 h_j^1 + \cdots + k_n h_j^n \equiv 0 \pmod{d_j} \quad \text{for every } j = 1, \dots, m \quad (\text{where } h^i = h_1^i \dots h_m^i). \quad (51)$$

This in turn holds iff there exists an integer vector $l = l_1 \cdots l_m$ such that

$$k_1 h_j^1 + \cdots + k_n h_j^n + d_j l_j = 0 \quad \text{for every } j. \quad (52)$$

Remark that (52) is a system of m linear diophantine equations in the $n + m$ unknowns (k, l) . Furthermore $n = O(\log \mathfrak{g}) = m$ and h_j^i and d_j are integers of size $O(\log \mathfrak{g})$. Using theorem 4, a generating set of solutions

$$(k^\alpha, l^\alpha), \quad \alpha = 1, \dots, r \quad \text{for some } r = \text{polylog}(\mathfrak{g}) \quad (53)$$

can thus be computed in $\text{polylog}(\mathfrak{g})$ time. But then every solution k to (51) is generated by the elements k^α , implying that the latter form a generating set of D .

As argued above, it now follows that \mathcal{D} is generated by the operators $T(k^\alpha)$. Finally, given k^α the label of $T(k^\alpha)$ can be computed efficiently by repeatedly applying lemma 6, since $T(k^\alpha)$ is defined as a product of powers of the Pauli matrices σ_i . \square

We use shorthand notation $T(k^\alpha) =: V_\alpha$ for the generators of \mathcal{D} computed in lemma 9. Since V_α belongs to the diagonal subgroup, its label is $(c_\alpha, z^\alpha, 0)$ for some $c_\alpha \in \mathbf{Z}_{2\mathfrak{g}}$ and $z^\alpha \in G$, all of which are efficiently computable owing to lemma 9.

In summary so far, we have computed a generating set of \mathcal{D} and thus (via lemma 8) of every subgroup \mathcal{G}_x . Now let S denote the support of $|\psi\rangle$ and consider an arbitrary $x \in G$.

Theorem 6c then implies that $S = x + H$ if and only if $V_\alpha|x\rangle = |x\rangle$ for every α . This last condition holds iff

$$\chi_{z^\alpha}(x) = \gamma^{-c_\alpha} \quad \text{for every } \alpha = 1, \dots, r \quad (54)$$

where we have used that

$$V_\alpha|x\rangle = \gamma^{c_\alpha} Z(z^\alpha)|x\rangle = \gamma^{c_\alpha} \chi_{z^\alpha}(x)|x\rangle. \quad (55)$$

The equations (54) are guaranteed to be satisfied by *some* x since we know from corollary 1 that the support S coincides with some coset $x + H$. Since the l.h.s. of (54) are \mathfrak{g} -th roots of unity, these equations can only have a solution if γ^{-c_α} is also a \mathfrak{g} -th root of unity. Thus there must exist $s_\alpha \in \mathbf{Z}_\mathfrak{g}$ such that $\gamma^{-c_\alpha} = e^{2\pi i s_\alpha / \mathfrak{g}}$. Given c_α , the number s_α can easily be computed. In conclusion, we find that

$$S = x + H \quad \text{iff} \quad \chi_{z^\alpha}(x) = e^{2\pi i s_\alpha / \mathfrak{g}} \quad \text{for every } \alpha = 1, \dots, r. \quad (56)$$

Theorem 3b can now be used to compute a solution $x \equiv x^0$ efficiently.

9.5 Sampling from $H + x^0$

Having determined the coset $H + x^0$ coinciding with the support of $|\psi\rangle$, we show that a random element in this coset can be generated efficiently. Recalling corollary 2 this will complete the proof of theorem 1.

Lemma 10. *A random element of $H + x^0$ is generated in $\text{polylog}(\mathfrak{g})$ time as follows:*

- (a) Generate $t_i \in \mathbf{Z}_\mathfrak{g}$ uniformly at random, where $i = 1, \dots, n$.
- (b) Compute $h := \sum t_i h^i$
- (c) Output $h + x^0$.

Proof: first we show that the procedure (a)-(b) generates a uniformly random element of H . Remark that this procedure can only output elements within H and, since H is generated by the h^i , each element of this subgroup occurs with nonzero probability. We show that all these probabilities are equal; some care is required since the generators h^i are not guaranteed to be independent. The map $\nu : \mathbf{Z}_\mathfrak{g}^n \rightarrow H$ defined by $\nu(t) := \sum t_i h^i$ (where $t = t_1 \dots t_n$) is easily seen to be a homomorphism between the additive group $\mathbf{Z}_\mathfrak{g}^n$ and H . With (a)-(b) every $h \in H$ is generated with probability

$$p(h) = |\{t : \nu(t) = h\}| / \mathfrak{g}^n. \quad (57)$$

The kernel N of ν is a subgroup of $\mathbf{Z}_\mathfrak{g}^n$ since ν is a homomorphism. Let $t_h \in \mathbf{Z}_\mathfrak{g}^n$ be a single vector satisfying $\nu(t_h) = h$ and consider the coset $t_h + N$. We now claim that

$$t_h + N = \{t : \nu(t) = h\}. \quad (58)$$

The inclusion \subseteq holds trivially. To prove the converse inclusion, note that every t satisfying $\nu(t) = h$ can be written as $t = t_h + t - t_h$, where $t - t_h \in N$. This shows (58). Note further that each coset of N has the same cardinality as N . Therefore $p(h)$ is equal to $|N|/|\mathfrak{g}^n|$ and thus independent of h . This shows that this probability distribution is uniform over H . It easily follows that (a)-(b)-(c) generates a uniformly random element in $H + x^0$. Finally, it is easily verified that this procedure can be implemented efficiently. \square

10 Proof of theorem 2

Suppose that there exists a normalizer circuit \mathcal{C} satisfying $\|\mathcal{C} - U_{\text{me}}\| < \delta$ with $\delta = 1 - 1/\sqrt{2}$. Our goal is to arrive at a contradiction.

Defining $F(x, y) := (x, y + a^x)$, the inequality $\|\mathcal{C} - U_{\text{me}}\| < \delta$ implies that $\|\mathcal{C}|x, y\rangle - |F(x, y)\rangle\| < \delta$ and thus

$$|\langle F(x, y) | \mathcal{C} | x, y \rangle| > 1 - \delta = 1/\sqrt{2} \quad (59)$$

for every x and y . Since $\mathcal{C}|x, y\rangle$ is a coset state, owing corollary 1 this state must be a uniform superposition. If this state is a uniform superposition of two or more standard basis states, then $|\langle u, v | \mathcal{C} | x, y \rangle| \leq 1/\sqrt{2}$ for all standard basis states $|u, v\rangle$, yielding a contradiction with (59). We therefore conclude that $\mathcal{C}|x, y\rangle$ is just a single basis state. In particular (59) implies that $\mathcal{C}|x, y\rangle$ must then coincide with $|F(x, y)\rangle$ for every x and y and thus $\mathcal{C} = U_{\text{me}}$. We now use the following lemma, to be proved at the end of this section:

Lemma 11. *Consider an arbitrary finite Abelian group G . Consider a unitary gate $U : |g\rangle \rightarrow |F(g)\rangle$, where F is a permutation of G , which can be realized as a normalizer circuit. Then there exists an automorphism α and $t \in G$ such that $F(g) = \alpha(g) + t$ and F is called affine.*

Since we have showed that $\mathcal{C} = U_{\text{me}}$, lemma 11 implies that the exponential function $F(x, y) = (x, y + a^x)$ must be an affine function on the group $G = \mathbf{Z}_{2^m} \times \mathbf{Z}_N$. We show that this is not the case, implying that our initial assumption about the existence of \mathcal{C} was false. Suppose that there exists an automorphism α of G with matrix representation A and $t \in G$ such that $F(x, y) = \alpha(x, y) + t$. By evaluating this identity in $(0, 0)$, $(1, 0)$ and $(0, 1)$ is easy to show that then $t = (0, 1)$ and

$$A = \begin{bmatrix} 1 & 0 \\ a - 1 & 1 \end{bmatrix}. \quad (60)$$

The identity $F(x, y) = \alpha(x, y) + t$ then implies that

$$a^x \equiv (a - 1)x + 1 \pmod{N} \quad \text{for all } x \in \{0, \dots, 2^m - 1\}. \quad (61)$$

Considering $x = 2$ yields $(a - 1)^2 \equiv 0 \pmod{N}$. Thus, since $a \neq 1$, there exists a nonzero integer k such that $(a - 1)^2 = kN$. Since $N = pq$ with $p \neq q$ it follows that both p and q divide $(a - 1)^2$. But then p and q also divide $a - 1$. In other words, both p and q occur in the prime decomposition of $a - 1$. It follows that $N = pq$ divides $a - 1$. This leads to a contradiction since $a \leq N$. This proves theorem 2.

We conclude by proving lemma 11. Suppose that U can be realized as a normalizer circuit. Then for every $g \in G$ there exists $\alpha(g) \in G$ such that $UX(g)U^\dagger = X(\alpha(g))$. Indeed since U acts as a permutation on the standard basis and since $X(g)$ is also a permutation, $UX(g)U^\dagger$ is again a permutation; since by assumption U maps $X(g)$ onto a Pauli operator and since only the Pauli operators $X(h)$ are permutations, the claim follows. Using that

$$UX(g)U^\dagger UX(g')U^\dagger = UX(g + g')U^\dagger \quad (62)$$

one shows that α is an automorphism. Applying both sides of the identity $UX(g)U^\dagger = X(\alpha(g))$ to $|0\rangle$ yields

$$F(g + F^{-1}(0)) = \alpha(g) \text{ for every } g \quad (63)$$

Setting $h := g + F^{-1}(0)$ and $t := -\alpha(F^{-1}(0))$ and using that α is an automorphism yields $F(h) = \alpha(h) + t$ for every $h \in G$. This proves lemma 11.

11 Further examples of normalizer gates

In this final section we give further illustrations of group automorphisms and quadratic functions, and the associated normalizer gates. See also section 2.

We first consider automorphisms. Consider an arbitrary $a \in \mathbf{Z}_d$. Then the multiplication map $M_a : x \rightarrow ax$ is an endomorphism of the cyclic group \mathbf{Z}_d . If a is coprime to d then a has a multiplicative inverse i.e. there exists b in \mathbf{Z}_d such that $ab = 1$ [15]. Therefore, for such a the multiplication map is an automorphism. Second, consider the group $\mathbf{Z}_{d_1} \times \mathbf{Z}_{d_2}$, fix $c \in \mathbf{Z}_{d_2}$ satisfying $d_1c \equiv 0 \pmod{d_2}$ and define [27]:

$$F_c : (x, y) \in \mathbf{Z}_{d_1} \times \mathbf{Z}_{d_2} \rightarrow (x, y + xc). \quad (64)$$

Owing to lemma 3, the map F_c is an endomorphism. Furthermore this map has an inverse viz. F_{-c} , showing that F_c is an automorphism. Remark that for $d_1 = d_2 = 2$ and $c = 1$ one obtains the CNOT gate. Considering an arbitrary Abelian group (1), applying M_a to any cyclic group \mathbf{Z}_{d_i} (with a coprime to d_i) and applying F_c to any pair $\mathbf{Z}_{d_i} \times \mathbf{Z}_{d_j}$ (with $d_i c \equiv 0 \pmod{d_j}$) yield automorphisms of G . Finally, since the set of all automorphisms of any group is a group as well, further examples can be obtained by taking products and inverses (and combinations thereof) of the examples given above.

Next we give examples of quadratic functions. Consider $a \in \mathbf{Z}_d$. One can straightforwardly verify that the functions

$$x \in \mathbf{Z}_d \rightarrow e^{2\pi i ax/d} \quad \text{and} \quad x \in \mathbf{Z}_d \rightarrow e^{2\pi i ax^2/d} \quad (65)$$

are quadratic functions for \mathbf{Z}_d . Remark that (65) coincides with the character χ_a . Second, consider $\mathbf{Z}_{d_1} \times \mathbf{Z}_{d_2}$ and $c \in \mathbf{Z}_{d_2}$ satisfying $d_1c \equiv 0 \pmod{d_2}$. Then

$$(x, y) \in \mathbf{Z}_{d_1} \times \mathbf{Z}_{d_2} \rightarrow e^{2\pi i cxy/d_2} \quad (66)$$

is quadratic. This can be shown by using the following lemma:

Lemma 12. *Consider an arbitrary finite Abelian group. For every endomorphism ϖ , the following function is quadratic: $f_\varpi(g) := \chi_g(\varpi(g))$.*

Proof: the proof is obtained straightforwardly by using basic properties of χ_g and the fact that ϖ is an endomorphism. \square

Using lemma 3 one notes that $\varpi : (x, y) \rightarrow (0, cx)$ is an endomorphism. The associated function f_ϖ is precisely (66), showing that the latter is indeed quadratic. Remark that for $d_1 = d_2 = 2$ and $c = 1$ the function (66) maps the pair of bits (x, y) to $(-1)^{xy}$. The associated quadratic phase gate is the CZ gate, cf. section 2. Lemma 12 allows to generate an entire class of quadratic functions. Interestingly, not every quadratic function is of the

form f_{ϖ} . To see this, consider again a single cyclic group and denote $q_a(x) := ax(x+d)$ where $a, x \in \mathbf{Z}_d$ but $q_a(x)$ is computed *over the integers* (i.e. not within \mathbf{Z}_d). Then

$$x \rightarrow e^{\pi i q_a(x)/d} \quad (67)$$

is a quadratic as well; the proof is straightforward but somewhat tedious and we refer to appendix D. For $d = 2$ and $a = 1$ one simply has $x \rightarrow (-i)^x$ which gives rise to the single-qubit phase gate $\text{diag}(1, -i)$ as discussed in section 2. Note that the function (67) cannot be of the form f_{ϖ} . Indeed the function values of the former are $2\mathbf{g}$ -th roots of unity, whereas the function values of the latter are \mathbf{g} -th roots of unity.

Applying (65) and (67) to any single cyclic subgroup of a general Abelian group G and applying (66) to any pair of cyclic subgroups yield quadratic functions of G . Finally we remark that the set of a quadratic functions of G can be shown to be a group under the multiplication of functions; the neutral element is the trivial function $g \rightarrow 1$ and the inverse of ξ is its complex conjugate. Thus further examples of quadratic functions (and the associated quadratic phase gates) can be obtained by taking products and complex conjugates, and combinations thereof, of the examples given above.

Acknowledgements

I thank J. Bermejo-Vega for discussions, W. Dür for suggestions to improve the presentation of the manuscript, and A. Schrijver, A. Schönhage and A. Storjohann for providing references on linear diophantine equations.

A Efficient quantum circuits

We show that every automorphism gate and quadratic phase gate can be realized as an efficient quantum circuit.

Let α be an automorphism of G , specified in terms of its matrix representation. Recall that the function α is invertible and that $\alpha(g)$ is efficiently computable on input of g (cf. section 6). Thus there exists an efficient circuit of classical reversible gates that computes $g \rightarrow \alpha(g)$. But then the standard translation of these classical gates into unitary operators yield an efficient quantum circuit to compute $|g\rangle \rightarrow |\alpha(g)\rangle$.

Let ξ be a quadratic function, specified in terms of its standard representation as in section 2. Recall from section 6 that $g \rightarrow \xi(g)$ can be computed efficiently classically. Furthermore each $\xi(g)$ is a $2\mathbf{g}$ -th root of unity. Writing $\xi(g) = \exp[i\pi n(g)/\mathbf{g}]$ where $n(g)$ is an integer between 0 and $2\mathbf{g}$, we denote by $f(g)$ the bit string corresponding to the binary expansion of $n(g)$. Note that each $f(g)$ has length $k = O(\log \mathbf{g})$ and that $g \rightarrow f(g)$ can be computed efficiently classically. We will write $f(g) = f_1(g) \cdots f_k(g)$ where each $f_i(g)$ is a bit; the relation to $n(g)$ is given by $n(g) = \sum_{j=1}^k f_j(g) 2^{j-1}$. It follows that

$$\xi(g) = \prod \omega_j^{f_j(g)}; \quad \omega := \exp[i\pi 2^{j-1}/\mathbf{g}]. \quad (68)$$

Now consider the following procedure: start with the input $|g\rangle|0\rangle$ where $|0\rangle$ represents a string of k ancillary qubits in the state $|0\rangle$. Since $f(g)$ can be computed efficiently classically, there exist a efficient quantum circuit realizing $U_f : |g\rangle|a\rangle \rightarrow |g\rangle|f(g) + a\rangle$ for every k -bit string a (where $f(g) + a$ is computed modulo 2) so that in particular $|g\rangle|0\rangle$ is mapped to

$|g\rangle|f(g)\rangle$. We now apply the single-qubit gate $\text{diag}(1, \omega_j)$ to qubit j in the second register, for all j . Using (68) this maps

$$|g\rangle|f(g)\rangle \rightarrow \xi(g)|g\rangle|f(g)\rangle. \quad (69)$$

Finally, applying the inverse of U_f maps $\xi(g)|g\rangle|f(g)\rangle \rightarrow \xi(g)|g\rangle|0\rangle$. This yields the desired efficient circuit that realizes U_ξ .

B Proof of theorem 3

B.1 Proof of theorem 3a

Let h^1, \dots, h^n represent a generating set of H . The orthogonal subgroup H^\perp consists of all $g \in G$ satisfying $\chi_{h^i}(g) = 1$ for all i . Let $U(1)$ denote the (multiplicative) Abelian group of all complex phases, where 1 is the neutral element. Let $U(1)^n$ be the group of all n -tuples of complex phases where the group operation is componentwise multiplication; the neutral element is $(1, \dots, 1)$. For every $g \in G$ we extend the domain of the function χ_g to the set of all m -tuples of integers i.e. if $k = k_1 \cdots k_m \in \mathbf{Z}^m$ with k_i integer, one has

$$\chi_g : k \rightarrow \exp 2\pi i \left(\frac{g_1 k_1}{d_1} + \cdots + \frac{g_m k_m}{d_m} \right), \quad (70)$$

where, with a slight abuse of notation, we denote the function (70) by χ_g as well. We now introduce the following maps.

$$\begin{aligned} \eta : G &\rightarrow U(1)^n \text{ maps } g \rightarrow (\chi_{h^1}(g), \dots, \chi_{h^n}(g)) \\ \varphi : \mathbf{Z}^m &\rightarrow U(1)^n \text{ maps } k \rightarrow (\chi_{h^1}(k), \dots, \chi_{h^n}(k)). \\ \pi : \mathbf{Z}^m &\rightarrow G \text{ maps } k \rightarrow (k_1 \bmod d_1, \dots, k_m \bmod d_m). \end{aligned}$$

Remark that π is the natural “projection” of \mathbf{Z}^m onto G . Every $g \in G$ can naturally be interpreted as an m -tuple of integers; we denote this tuple by \tilde{g} if we want to emphasize this way of interpreting g . Note that $\pi(\tilde{g}) = g$.

Lemma 13. *The maps η, φ and π are group homomorphisms. Furthermore $\varphi = \eta \cdot \pi$. Finally, H^\perp coincides with the kernel of η .*

Proof: all properties can be verified by a straightforward application of definitions, except the identity $\varphi = \eta \cdot \pi$ which we address next. Let $k = k_1 \cdots k_m$ be any m -tuple of integers and denote $\pi(k) = l_1 \cdots l_m$ with $l_i \in \mathbf{Z}_{d_i}$. Then (by the definition of $x \bmod y$) there exist integers c_i such that $k_i = l_i + c_i d_i$. Insertion in (70) shows that $\chi_g(k) = \chi_g(\pi(k))$ for every $g \in G$. It follows that $\varphi(k) = \eta(\pi(k))$. \square

To prove theorem 3 our goal is to compute a generating set of the kernel of η . Our approach will be to first compute a generating set of the kernel of φ , and then to map this solution to a generating set of $\ker \eta$ via the projection π . Remark that $k \in \ker \varphi$ iff

$$a_{i1}k_1 + \cdots + a_{im}k_m \equiv 0 \pmod{\mathfrak{g}}; \quad \text{where } a_{ij} := \mathfrak{g}h_j^i/d_j. \quad (71)$$

Remark that d_i divides \mathfrak{g} so that the a_{ij} are integers. Equivalent to (71), there exists a tuple of integers $s = s_1 \dots s_n$ such that $\sum a_{ij}k_j + \mathfrak{g}s_i = 0$. This is a system of linear diophantine equations in the unknowns (k, s) . Owing to theorem 4 we can find a generating

set of solutions $(k^1, s^1), \dots, (k^r, s^r)$ in $\text{polylog}(\mathfrak{g})$ time. It then follows that the k^i form a generating set of solutions to the equations (71) and thus are a generating set of $\ker \varphi$.

Define $g^i := \pi(k^i)$ for every i . Note that each g^i can be computed efficiently since this computation involves taking m remainders. We now claim that the g^i generate $\ker \eta$. First, we have $g^i \in \ker \eta$: indeed by construction $k^i \in \ker \varphi$ and owing to lemma 13 we have $\varphi = \eta \cdot \pi$, implying that $\pi(k^i) \in \ker \eta$. Second, we show that every element in the kernel of η can be written as an integer linear combination of the g^i . Consider an arbitrary $g \in \ker \eta$. Interpreting g as an element in \mathbf{Z}^m we thus have $\tilde{g} \in \ker \varphi$. Since the k^i generate $\ker \varphi$, there exist integers c_i such that $\tilde{g} = \sum c_i k^i$. Since π is a homomorphism, it follows that $\pi(\tilde{g}) = \sum c_i \pi(k^i)$. Using that $\pi(\tilde{g}) = g$ and $g^i = \pi(k^i)$ we conclude that $g = \sum c_i g^i$ as desired. This proves theorem 3a.

B.2 Proof of theorem 3b

The proof method is analogous to above (although the argument is simpler since we do not need to compute a generating set of solutions but merely a single solution) and we just give a sketch. Using the notations introduced above, our goal is to compute a $g \in G$ satisfying

$$\eta(g) = (e^{2\pi i s_1/\mathfrak{g}}, \dots, e^{2\pi i s_n/\mathfrak{g}}) =: v. \quad (72)$$

Analogous to above (see (71)) we can write down a system of diophantine equations which has a solution iff there exists $k \in \mathbf{Z}^m$ satisfying $\varphi(k) = v$. Deciding whether a solution exists can be done efficiently by virtue of theorem 4. If a solution exists, owing to the same theorem such a solution k can be computed efficiently. We then compute $g := \pi(k)$ which can also be done efficiently. Since $\varphi = \eta \cdot \pi$ it follows that $\eta(g) = v$ as desired.

B.3 Proof of theorem 3c

Lemma 14. *Let α be an automorphism of G with matrix representation A . Then there exists a unique matrix B where each element in the i -th row of B belongs to \mathbf{Z}_{d_i} and such that the following identities hold:*

$$d_j B_{ij} \equiv 0 \pmod{d_i} \quad \text{for every } i, j \quad (73)$$

$$\sum_{j=1}^m B_{ij} A_{ji} \equiv 1 \pmod{d_i} \quad \text{for every } i \quad (74)$$

$$\sum_{j=1}^m B_{kj} A_{ji} \equiv 0 \pmod{d_k} \quad \text{for every } k \neq i. \quad (75)$$

Furthermore this unique solution B is the matrix representation of α^{-1} .

Proof: to show existence, it is straightforward to verify (with lemma 3) the matrix representation of α^{-1} satisfies (73)-(75). To prove uniqueness, (73) and lemma 3 imply that there exists an endomorphism β of G with matrix representation B . Conditions (74)-(75) imply that $\beta\alpha(e^i) = e^i$ for every i . Since α and β are endomorphisms, it follows that $\beta\alpha(g) = g$ for every $g \in G$. But then β must be identical to the inverse of α , which always exists and which is unique. But then also B is unique. \square

Our proof of theorem 3c now proceeds as follows. For every $m \times m$ integer matrix X , let $\Pi(X)$ denote the matrix obtained by replacing the entry X_{ij} by $X_{ij} \bmod d_i$. We now

first compute an integer matrix B satisfying (73)-(75), thus for the moment ignoring the constraint that the i -th row of B should only contain elements of \mathbf{Z}_{d_i} . Computing such an integer matrix solution can be done efficiently by reducing the problem to a system of diophantine equations (by introducing extra integers variables as before). Given an integer solution B , we compute $B' := \Pi(B)$; this can be done efficiently. Remark that every element in the i -th row of B' does belong to \mathbf{Z}_{d_i} . We now claim that B' is still a solution to (73)-(75). By definition of the map Π , for every (i, j) there exists an integer C_{ij} which is an integer multiple of d_i , such that $B'_{ij} = B_{ij} + C_{ij}$. Since each C_{ij} is an integer multiple of d_i , one has

$$\begin{aligned} d_j C_{ij} &\equiv 0 \pmod{d_i} \quad \text{for every } i, j \\ \sum_{j=1}^m C_{kj} A_{ji} &\equiv 0 \pmod{d_k} \quad \text{for every } k. \end{aligned} \tag{76}$$

Using (76) in combination with the property that B solves (73)-(75), it follows that B' is a solution to these equations as well. Thus B' is the matrix representation of α^{-1} .

C Proof of theorem 6

The normalized sum $\rho := (1/|\mathcal{G}|) \sum U$ over all $U \in \mathcal{G}$ is the orthogonal projector on the space of all +1 common eigenvectors of \mathcal{G} (this property holds for arbitrary unitary groups, see e.g. [24]). Since $|\psi\rangle$ is the unique joint eigenvector with eigenvalue 1 by assumption, we thus have $\rho = |\psi\rangle\langle\psi|$. Consider an arbitrary g such that $\langle g|\psi\rangle \neq 0$. Then

$$|\psi\rangle \propto |\psi\rangle\langle\psi|g\rangle = \frac{1}{|\mathcal{G}|} \sum U|g\rangle. \tag{77}$$

This shows that all amplitudes $\langle h|\psi\rangle$ are zero whenever h lies *outside* the orbit O_g . Now consider an arbitrary $h \in O_g$. Then, by definition of O_g , there exist $U \in \mathcal{G}$ satisfying $U|g\rangle = d_g|h\rangle$ where d_g is the diagonal entry of D on position g . Remark that d_g is a complex phase since U is unitary. Using that $U^\dagger|\psi\rangle = |\psi\rangle$ it follows that $\langle h|\psi\rangle = d_g\langle g|\psi\rangle$. This proves that all amplitudes $\langle h|\psi\rangle$ with $h \in O_g$ are equal in modulus. This shows (a) and (b).

Next we prove (c). Suppose first that $O_g = S$ and consider a generator V_i . Then there exists a complex phase ξ such that $V_i|g\rangle = \xi|g\rangle$. It follows that

$$\langle\psi|V_i|g\rangle = \xi\langle\psi|g\rangle. \tag{78}$$

Using that $\langle\psi|V_i = \langle\psi|$ since V_i belongs to \mathcal{G} , one finds that $\langle\psi|g\rangle = \xi\langle\psi|g\rangle$. Since $O_g = S$ and $g \in O_g$, it follows that $\langle\psi|g\rangle \neq 0$. With (78) this shows that $\xi = 1$. Conversely, suppose that g satisfies $V_i|g\rangle = |g\rangle$ for every i . As a consequence, $V|g\rangle = |g\rangle$ for every $V \in \mathcal{G}_g$. Using that $\rho = |\psi\rangle\langle\psi|$ yields

$$|\langle\psi|g\rangle|^2 = \frac{1}{|\mathcal{G}|} \sum_{U \in \mathcal{G}} \langle g|U|g\rangle = \frac{1}{|\mathcal{G}|} \sum_{U \in \mathcal{G}_g} \langle g|U|g\rangle = \frac{|\mathcal{G}_g|}{|\mathcal{G}|} \neq 0. \tag{79}$$

In the second identity we have used that $\langle g|U|g\rangle = 0$ if $U \notin \mathcal{G}_g$ since U is monomial; in the third identity we used that $U|g\rangle = |g\rangle$ for every $U \in \mathcal{G}_g$. Finally we used that $|\mathcal{G}_g| \neq 0$ since $I \in \mathcal{G}_g$.

D The quadratic functions (67)

We prove that the function (67), denoted here by ξ , is quadratic. Consider $x, y \in \mathbf{Z}_d$. Within the present section, we denote by $x + y$ the sum of these elements over the integers, whereas $x \oplus y$ denotes the sum in \mathbf{Z}_d . With this notation, one has

$$x \oplus y = x + y \bmod d \quad \text{and} \quad q_a(x) = ax(x + d) \quad (80)$$

It follows from the left-hand equation in (80) that there exists an integer b such that $x \oplus y = x + y + bd$. Consequently,

$$q_a(x \oplus y) = q_a(x + y + qd) = q_a(x) + q_a(y) + 2axy + ac \quad (81)$$

$$\text{where } c := 2xbd + 2ybd + d^2b(b + 1). \quad (82)$$

Remark that c is an integer multiple of $2d$. Indeed, both terms $2xbd$ and $2ybd$ are obviously multiples of $2d$; furthermore $b(b + 1)$ is even for every integer b , so that $d^2b(b + 1)$ is also a multiple of $2d$. It follows that $\exp[i\pi ac/d] = 1$ and thus with (81) one finds

$$\begin{aligned} \xi(x \oplus y) &= e^{i\pi[q_a(x) + q_a(y) + 2axy + ac]/d} \\ &= e^{i\pi q_a(x)/d} e^{i\pi q_a(y)/d} e^{2\pi i axy/d} e^{\pi i ac/d} = \xi(x)\xi(y)B(x, y), \end{aligned} \quad (83)$$

with $B(x, y) := \exp[2\pi i axy/d]$. Since the function B is bilinear, this shows that ξ is quadratic.

References

- [1] R. Jozsa, Proc. R. Soc. Lond. A 8 454 pp. 323-337 (1998).
- [2] A. Childs and W. Van Dam, Rev. Mod. Phys. 82, pp. 1-52 (2010).
- [3] P. W. Shor, , SIAM J. Comp 26(5), pp. 1484-1509 (1997).
- [4] D. Boneh and R. J. Lipton, In Don Coppersmith, Ed., CRYPTO '95, Lecture Notes in Computer Science, pages 424-437. Springer-Verlag (1995).
- [5] C. Lomont, quant-ph/0411037 (2004).
- [6] D. Gottesman, talk at International Conference on Group Theoretic Methods in Physics (1998), see also arXiv:quant-ph/9807006.
- [7] D. Aharonov, Z. Landau and J. Makowsky, arXiv:quant-ph/0611156 (2006).
- [8] N. Yoran and A. Short, Phys. Rev. A 76, 042321 (2007).
- [9] D. E. Browne, New J. Phys. 9 (146) (2007).
- [10] N. Yoran and A. J. Short, Phys. Rev. A 76, 060302(R) (2007)
- [11] I. Markov and Y. Shi, SIAM J. Comp. 38 (3) 963-981 (2008).
- [12] N. Yoran, arXiv:0802.1156.

- [13] A. Yu. Kitaev, quant-ph/9511026 (1995).
- [14] M. Mosca and C. Zalka, Int. J. Quant. Inf. 2(1), 91100 (2004).
- [15] H. Hasse, *Number Theory*, Berlin, Springer-Verlag (1980).
- [16] R. Jozsa, Computing in Science and Engineering 3(2), 34-43 (2001).
- [17] W. Fulton and J. and Harris, *Representation Theory - A First Course*, Springer Verlag (1991).
- [18] D. Gottesman, Chaos Solitons Fractals 10:1749-1758 (1999).
- [19] E. Hostens, J. Dehaene and B. De Moor, Phys. Rev. A 71, 042315 (2005).
- [20] W. A. Blankinship, Communications of the ACM 9(7) 514 (1966).
- [21] J. von zur Gathen and M. Sieveking, *Weitere zum Erfüllungsproblem polynomial äquivalente kombinatorische Aufgaben*. In Komplexität von Entscheidungsproblemen, Lecture Notes in Computer Science 42, 49-71, Springer-Verlag (1976).
- [22] R. Kannan and A. Bachem, SIAM J. Computing, 8 499507 (1979).
- [23] T.-W. Chou and G. E. Collins, SIAM J. Comp., 11(4):687708 (1982).
- [24] M. Van den Nest, New J. Phys. 13 123004 (2011).
- [25] Since $\mathfrak{g} = d_1 \cdots d_m$, one has $\log_2 \mathfrak{g} = \sum_{i=1}^m \log_2 d_i$. Since $d_i \geq 2$ one has $\log_2 d_i \geq 1$. This implies that $\log_2 \mathfrak{g} \geq m$.
- [26] Consider $G = \mathbf{Z}_2^m$. For every bit string $a = a_1 \cdots a_m \in G$ one finds $X(a) = X^{a_1} \otimes \cdots \otimes X^{a_m}$ and $Z(a) = Z^{a_1} \otimes \cdots \otimes Z^{a_m}$, where X and Z represent the standard σ_x and σ_z Pauli matrices, respectively.
- [27] Here cx is the element in \mathbf{Z}_{d_2} given by $c + \cdots + c$ with x summands.